DOCUMENT RESUME

ED 050 138 TM 000 525

AUTHOR Flanagan, John C.
TITLE A Critique of the Measurement and Instrumentation Aspects of Educational Evaluation and Decision-Making.
INSTITUTION American Institutes for Research, Palo Alto, Calif.
PUB DATE Feb 71
NOTE 5p.; Paper presented at the Annual Meeting of the American Educational Research Association, Symposium on "Critique of the Report of the Phi Delta Kappa Study Committee on Evaluation," New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Data Collection, *Decision Making, Educational Objectives, *Evaluation, *Evaluation Methods, Evaluation Techniques, Formative Evaluation, *Individualized Programs, *Measurement Techniques, Student Participation, Summative Evaluation
IDENTIFIERS *Phi Delta Kappa Study Committee on Evaluation

ABSTRACT

This critique of the Phi Delta Kappa Study Committee on Evaluation's report opens with a short description of the report; its definition of evaluation as "the process of delineating, obtaining, and providing useful information for judging decision alternatives," its detailed description of the decision-making process together with descriptions of possible decision settings, decision types, and problems related to decision-making. Four types of evaluation referred to in the PDK report are discussed: context, input, process, and product. Support is given to the report's emphasis on context evaluation and its division into contingency and congruence modes. The distinction made between context and product evaluations is endorsed, but the proposed development of specifications and procedures for data collection is considered inadequate. The suggestion to use individual students in evaluation studies is regarded as relevant and valuable but new data collection techniques are recommended. A case is made for greater involvement of students in determining individual educational objectives, particularly in their later educational years. (LR)

A CRITIQUE OF THE MEASUREMENT AND INSTRUMENTATION ASPECTS OF EDUCATIONAL EVALUATION AND DECISION-MAKING

John C. Flanagan

American Institutes for Research

As one of a panel of five reactors to this report, I feel a little bit like one of the five blind men describing the elephant. Unlike the blind men, my colleagues and I all see this elephant, but our descriptions are likely to be rather dissimilar because of our special fields of interest and our varied previous experiences. Thus, the descriptions, although all based on the same 532-page elephant, can be expected to be quite different.

In my description the emphasis will be on techniques of measurement, data collection, and the central role of the individual student in evaluation activities. Before proceeding to specific points, some general impressions seem in order. The report is comprehensive, detailed, and analytical. It analyzes evaluation into stages occurring in various settings, having various scopes, and providing information relevant to various types of decisions.

The report is based on a specific definition of evaluation which is: "Educational evaluation is the process of delineating, obtaining, and providing useful information for judging decision alternatives."

This definition is followed by a discussion of four stages in the process of decision-making including seventeen specific elements. In addition to this study of the decision process, there is a detailed description of possible decision settings, decision models, types of decisions, and some problems related to decision-making.

Presented at the Symposium on "Critique of the Report of the Phi Delta Kappa Study Committee on Evaluation," AERA Meetings, New York City, February 1971.

In this chapter and the chapters which follow on criteria, values, information and systems theory, and evaluation methodology, the emphasis seems to be on delineating and discussing all possibilities rather than on the practical side of the conduct of educational evaluation.

In Chapter Seven, the four types of evaluation are presented together with a general model for conducting any one of these types of evaluation. The three steps which are proposed for all types of evaluation are delineating, obtaining, and providi

The four types of evaluation are:

1. Context evaluation has as its purpose "to provide a rationale for determinati of objectives. Specifically, it defines the relevant environment, describes the desired and actual conditions pertaining to that environment, identifies unmet needs and unused opportunities, and diagnoses the problems that prevent needs from being met and opportunities from being used. The diagnosis of problems provides an essential basis for developing objectives whose achieveme will result in program improvement."

The authors state that context evaluation is the most basic kind of evaluation. The authors divide context evaluation into two modes - contingency and congruence. In the contingency mode evaluation searches for opportunities to improve the system by changing the objectives. The congruence mode evaluates the extent to which intended objectives are achieved.

This reviewer strongly endorses the emphasis on this type of evaluation and the distinction between the two modes for studying the objectives of the system. The discussion, however, seems to lack sufficient emphasis on needs and opportunities with respect to individual students and, although there is an emphasis on broad exploratory probing, it appears desirable that there be more specific provision for unplanned outcomes and the achievement of unintended objectives as well as those intended for the system.

2. Input evaluation is intended to provide the basis for selecting a design to achieve program objectives. This involves the study of relevant capabilities, strategies for achieving objectives, and specific designs for implementing a proposed strategy.

The authors point out that "techniques for input evaluation are lacking in education." One available technique which appears applicable and is not discussed is the method of explicit rationales.

3. Process evaluation is intended "to provide periodic feedback to persons responsible for implementing plans and procedures." To a substantial degree, what these authors have included in process evaluation has come to be known as formative evaluation following the terminology of Michael Scriven. The objectives of process evaluation are: to monitor the implementation of the design, to provide information needed for planned decisions during the implementation phase, and to maintain a record of the extent to which the project is actually implemented as designed. This type of evaluation is clearly of great importance.

4. Product evaluation measures and interprets the extent to which objectives were achieved. The criteria which are measured to perform this evaluation are classified as either instrumental or consequential following Scriven's terminology. Instrumental criteria refer to what have been frequently called intermediate criteria. Consequential criteria are those usually called ultimate criteria.

The authors point out that "in the assessment of objectives relating to adoption, product evaluation and context evaluation ultimately merge in the measurement of the impact of the total change effort on the overall system. Context evaluation then takes on the systematic functions of monitoring the total system and the ad hoc product evaluation is terminated."

In a later section the authors state "product evaluation assesses attainments of change projects within a system, and context evaluation assesses the impact of the obtained change on the total system." This distinction between context and product evaluation seems to be a useful one. In their general discussion of the features of their evaluation model, the authors again emphasize the basic importance of context evaluation and the need for a much more comprehensive data base to perform this function. Unfortunately, they do not seem to go far enough in developing specifications and procedures for collecting this very important data base. Educational systems have continued to operate with very little attention at either the local or national level to the study of the needs of individual students. The authors of this study have inserted two or three paragraphs suggesting the use of the individual student as the unit of measure in evaluation studies. The remarks are relevant and valuable. It would be desirable if their implications were carried through more fully in the subsequent discussions of implementing evaluation programs.

The later chapters of the report on implementing and administering evaluation programs need to be supplemented by handbook materials on what data to collect to study the needs and opportunities of the total educational system especially as it relates to the individual student. Some of the procedures used in recent years to obtain such data include intensive case studies of students on a sampling basis; follow-up studies of recent graduates to determine the utility of the knowledge and abilities achieved in school; and intensive studies of adults in various roles and activities to determine the specific educational objectives which would have been most appropriate for them during their study programs in school.

There is probably no more important problem in education at the present time than determining the educational objectives for each individual student. It is believed that during the later educational years, much of the responsibility for these decisions should be given to the student. To prepare him for taking such responsibility it is

believed that one should start in the primary grades by giving students some responsibility in planning and carrying out their educational programs. This will necessarily be limited in the early years, but the ability to take responsibility requires much practice.

This will require that the student know the specific knowledges and abilities required for many adult roles and activities. He must also know something about the nature of learning and individual differences and be able to estimate the extent of effort required for him to achieve a specified level of proficiency with respect to various types of content or ability. To assist the student in formulating his long-range educational and occupational goals the behavioral scientist needs detailed and extensive studies of students both during and following their exposure to specific educational experiences which can be made available to current students as a basis for making their decisions. A minor point regarding the present report is that in the view of these authors such behavioral scientists are clearly functioning as evaluators in providing the basis for individual decisions, however their functions appear to be broadly those of the behavioral scientist and not specifically those usually considered as appropriate for an evaluator.

To sum up this review of educational evaluation and decision-making as presented by these authors, the first point to be noted is that the definition selected by these authors includes only one type of evaluation in education and therefore should not be thought of as the only function of evaluation methods in the educational field. There are many instances in which evaluative data are very desirable even though no decisions have been defined and no actions are anticipated. However, for purposes of decision oriented educational evaluation, the report has much to commend it. The efforts of the seven members of the Phi Delta Kappa Commission on Evaluation represent an important step forward in increasing our understanding and ability to conduct effective educational evaluation studies. As the authors point out, this is only the beginning of an important effort to improve our educational programs.